

UNIVERSIDAD CARLOS III DE MADRID



INGENIERÍA INDUSTRIAL

PROYECTO FIN DE CARRERA

INFLUENCIA DEL MES DE NACIMIENTO EN EL FUTBOL

Autor: Alberto López García.

Tutor: Juan De Dios Tena Horrillo.

MADRID, 2015

En estudios anteriores, tras analizar los meses de nacimiento de los futbolistas, se dieron cuenta de que la mayoría de los jugadores habían nacido a principios de año. Se sabe que el éxito depende de la capacidad y del esfuerzo, sin embargo, el mes de nacimiento tiene una gran importancia. ¿Cómo es posible que haya un número mucho mayor de jugadores de fútbol de las ligas profesionales nacidos a primeros de año? ¿A qué se debe este efecto, denominado en trabajos anteriores como efecto de la edad relativa? ¿Por qué una persona nacida a primeros de año tiene una mayor probabilidad de convertirse en futbolista profesional? ¿Siempre ha existido este efecto de la edad relativa o es algo nuevo y hasta qué punto influye en las probabilidades de convertirse en jugador de fútbol? ¿Se presenta también este efecto a la hora de convertirse en entrenador?

Son muchas las cuestiones a resolver. Es conocido que en los seres vivos pequeñas causas producen grandes efectos. Una pequeña vibración en el agua puede provocar una salida repentina de un cocodrilo. Una música de fondo puede provocar un aumento de las ventas. Una llamada de emergencias puede provocar la movilización de decenas de vehículos incluso cuando la llamada sea falsa. Una pequeña causa puede provocar que la mayoría de los jugadores de las últimas temporadas nazcan a primeros de año. Sin duda el efecto de la edad relativa es enorme tal y como se apreciara en las gráficas, sin embargo, las causas que lo producen son muy pequeñas, incluso para muchos siguen siendo una incógnita. El efecto es evidente, la causa que lo provoca no. Esta relación: causa pequeña- efecto grande propia de los seres vivos es lo que da lugar al misterio y al

tedioso trabajo de investigación yendo hacia atrás en el tiempo para descubrir que pequeña causa provoco ese efecto. Aquí tratare de analizar el efecto de la edad relativa e intentare arrojar luz sobre la causa que está detrás de ella. Sin duda el efecto esta provocado por la causa. Una vez descubierta la causa, se puede eliminar el favoritismo hacia los jugadores nacidos a primeros de año, de la misma manera, que la eliminación de la música en las tiendas de ropa haría regresar el número de ventas a una situación normal.

AGRADECIMIENTOS

Quería dar las gracias a mis padres por haberme apoyado, a mis compañeros y amigos de la universidad con quien he pasado muy buenos momentos y al tutor de mi proyecto por la gran ayuda recibida para la realización del mismo.

ÍNDICE

1. INTRODUCCIÓN. SELECCIÓN EN EL FÚTBOL.

2. ANÁLISIS DEL MES DE NACIMIENTO DE LOS JUGADORES DE FUTBOL DURANTE LAS ÚLTIMAS TEMPORADAS.

3. ANÁLISIS DEL MES DE NACIMIENTO DE LOS JUGADORES DE FÚTBOL DESDE LA TEMPORADA 1928-1929 HASTA LA TEMPORADA 2013-2014.

3.1. Nacidos en España.

3.2. Comparativa por hemisferio.

3.3 Evolución temporal.

3.3.1 Conclusiones.

3.4 Entrenadores.

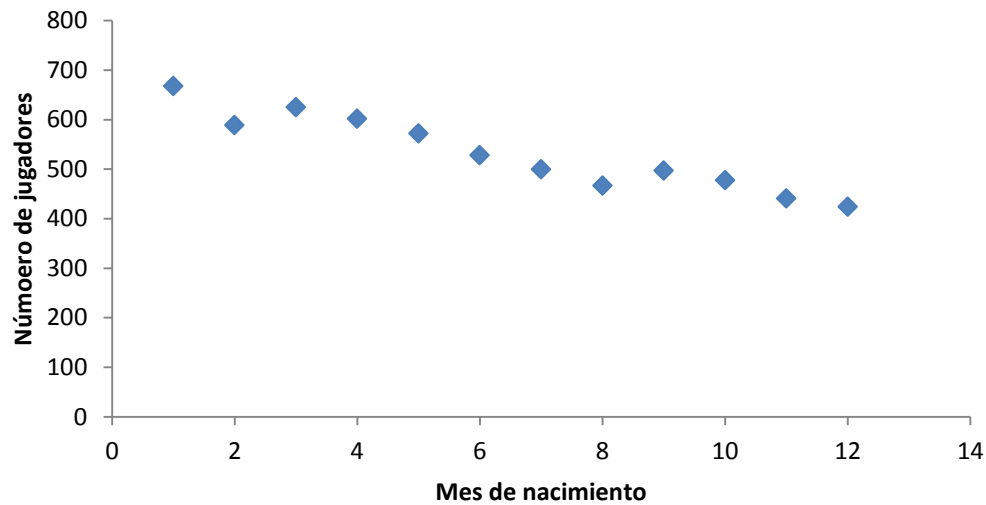
4. CONCLUSIONES.

1. INTRODUCCIÓN.SELECCIÓN DE LOS JOVENES DEPORTISTAS EN EL FÚTBOL.

La selección temprana de jóvenes talentos ha ido ganando importancia en los distintos deportes, particularmente en los que mueven mucho dinero como el fútbol. En el fútbol, los clubs de primera división cuentan en su mayoría con equipos de categorías inferiores que constituyen la materia prima de su producto final y de la economía de los clubs. Es evidente la importancia de una acertada selección y de una excelente formación.

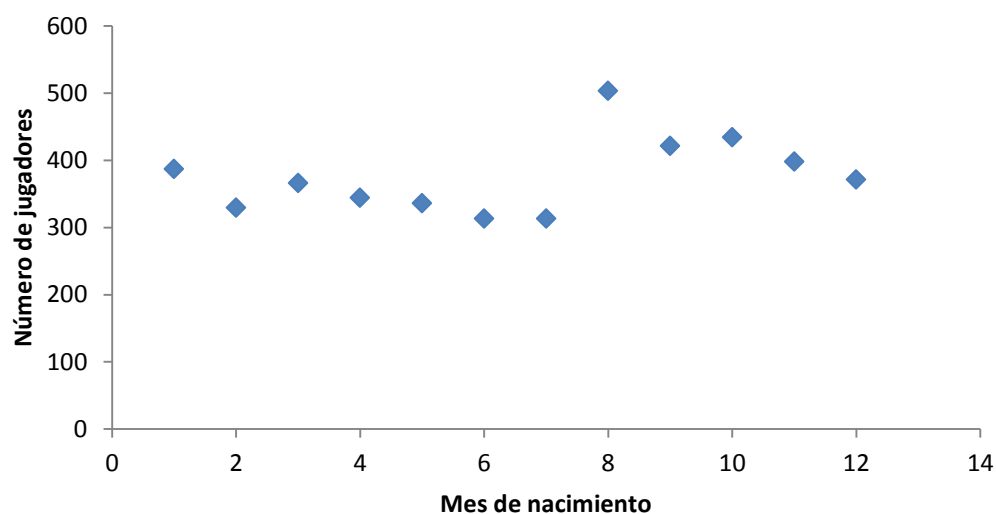
En estudios anteriores sobre deportes, se ha demostrado que existe un efecto de la edad relativa en múltiples ligas profesionales, el número de jugadores que llegan a primera división depende del mes de nacimiento de los jugadores. Un ejemplo es la liga canadiense de hockey NHL. La explicación de todo esto es bastante simple, en Canadá la fecha de corte para seleccionar jugadores de hockey en un grupo de edad es el 1 de enero. Así, un muchacho que cumpla diez años el 4 de enero podría estar jugando con alguien que no cumple años los diez años hasta finales de año; y a esa edad, en la preadolescencia, doce meses más o menos puede significar una enorme diferencia de madurez física. Para más información consúltese el libro: Gladwell, M. (2008): Fuera de serie (outliers). Porque unas personas tienen éxito y otras no. Taurus.).

Figura 1.1 Jugadores de la liga NHL de Canada desde la temporada 1951-52 hasta 2009-10.



En el béisbol en USA, por ejemplo, el mes de corte para seleccionar jugadores es agosto. Por tanto se observa una tendencia decreciente de agosto a julio. Ahora agosto sería el mes de nacimiento con más ventaja y julio el de más desventaja. Fíjese en la figura 1.2. En Spira, G. (2008): “The boys of late Summer” se puede consultar los datos con los que se construye la figura1.2.

Figura 1.2. Jugadores de las ligas mayores de baseball nacidos a partir de 1950 en USA.



Estos son estudios que se realizaron anteriormente a este. La diferencia entre este trabajo y los trabajos anteriores es que no solo se verá si existe un efecto de la edad relativa en los jugadores de fútbol sino que se analizara cómo ha evolucionado el efecto de la edad relativa desde el año 1900 hasta la actualidad. También se verá si influye el efecto de la edad relativa al variar de país y se estudiara si existe también este efecto de la edad relativa en los entrenadores. Para ello comenzare comentando como es la selección de los futuros jugadores de fútbol de la liga profesional.

Iniciación.

Una explicación de este proceso de iniciación y selección en el futbol se puede encontrar en Llames, R. (1999): “Revista de psicología del deporte, volumen 8, número 1”.

El inicio del fútbol se produce desde el juego en las horas de recreo, como actividad lúdica y no como un asunto con perspectiva profesional. Podría decirse que el inicio del fútbol comienza a una edad muy temprana por pura diversión. Es solamente un juego más como puede ser el escondite. En esto el fútbol es distinto a otros deportes que se aprenden más tarde como el patinaje, el tenis o el pádel. Entre los 5 a 6 años y los 8 a 9 años, una vez que el niño compite en algún tipo de liga, se inicia el proceso de captación por parte de los clubs. Entonces, el momento de captación oficial por parte de los clubs lo podemos considerar entre 8 y 10 años en la categoría Benjamin-Alevin que es cuando empieza la competición oficial federada.

El proceso de selección comienza con los ojeadores. Consiste en entrenadores, jugadores... cuya misión es detectar a aquellos jugadores que sobresalen. Generalmente los criterios son subjetivos como “este jugador tiene mucha técnica” o “este jugador tiene un talento especial, es muy bueno” aunque cada vez les exigen más rellenar un informe donde valoren varios aspectos poniendo una calificación en cada uno de ellos. Hay dos

métodos, el primero consiste en realizar varios partidos amistosos internos donde estarán los ojeadores y el segundo método consiste en que los ojeadores visiten partidos externos. El segundo método da una calificación menos fiable de los jugadores, sin embargo, llega a un número de jugadores mucho mayor.

La selección en fútbol es continua. Hay una primera selección con 9/10 años para formar los equipos Benjamín/Alevín. Aquí se capta a los niños, como ya hemos mencionado anteriormente, de las ligas escolares o de las ligas de fútbol sala (mini-benjamín y benjamín). Los criterios que se manejan, aunque nada estructurados, subjetivos, y no medibles, normalmente están relacionados con aspectos técnico-tácticos (“está muy bien puesto en el campo”, “maneja muy bien el balón”, “tiene muy buenas maneras”). Este primer equipo se irá disgregando a medida que los jugadores vayan avanzando en las diferentes categorías.

Entre el final de cada temporada y el inicio de la siguiente, se realiza una revisión de las plantillas, y se reestructuran para la siguiente temporada, confeccionándolas nuevamente equipo a equipo. Por lo tanto la selección se revisa anualmente, en función del rendimiento que ese jugador aporta al equipo o al club. Se irá dando la baja a jugadores que no se consideran válidos para el equipo superior, y a su vez se irán incorporando nuevos jugadores fichados a través de alguno de los sistemas mencionados anteriormente (Revista de psicología del deporte, volumen 8, número 1, 1999).

En esta tabla tenemos las categorías establecidas por edad de los participantes, internacionalmente se denominan con el prefijo “sub”.

Categorías en España.

Senior	Absoluto
Sub-19	Juvenil de tercer año
Sub-18	Juvenil
Sub-17	
Sub-16	Cadete
Sub-15	
Sub-14	Infantil
Sub-13	
Sub-12	Alevín
Sub-11	
Sub-10	Benjamín
Sub-9	
Sub-8	Pre-Benjamín
Sub-7	

Se puede observar como la formación empieza muy pronto. Naturalmente para entrar a dicha formación hay un proceso de selección del que se ha hablado antes. En este proceso de selección hay variables que influyen para ser seleccionado como la agilidad, la resistencia física, la capacidad estratégica... ¿acaso estas variables son independientes de la edad?. La respuesta es clara, no. Las personas no son constantes en el tiempo, todas las capacidades dependen de la edad que se tenga. Como diría Anjelica Huston: “donde hay edad, hay evolución”. También parece claro que a edades pequeñas estas variables dependen más del tiempo que a edades más avanzadas.

La capacidad para el fútbol se puede escribir así:

Capacidad para el fútbol (t)=f (agilidad (t), resistencia (t),...)

Es obvio que en edades de 0 a 20 años la capacidad para el fútbol aumenta con el tiempo. Para una determinada edad, por ejemplo 7,65 años, tendríamos una población que seguiría una distribución normal con una determinada media.

Si lo hacemos para otra determinada edad, por ejemplo 7,94 años, tendría otra distribución normal cuya media sería mayor que la anterior. Es decir la distribución es normal para personas con una misma edad.

Si nos movemos en el tiempo esa normal iría aumentando en media, es decir, iría aumentando la capacidad de fútbol. ¿Qué es lo que ocurre? que no se usa una edad exacta. Una persona tiene 15 años o 27 años, lo que no sucede es agrupar a personas con 1,2567 años de edad, es decir, dentro de mi grupo al que voy a seleccionar hay diferentes edades y, por tanto, diferentes normales.

Una persona de 6,1 años puede ser igual de buena que una de 6,9 años, sin embargo para la de 6,9 esa capacidad de fútbol puede ser normal y para la de 6,1 no, se trataría de un talento para su edad, de algo fuera de lo normal.

Si, por ejemplo, selecciono en un partido de fútbol en el que juegan chavales de un mismo curso académico lo que ocurrirá es que, tras la selección, habrá un mayor número de jugadores que nacieron en la primera parte del año que los que nacieron a finales. Lo que se va a realizar en el proyecto es un análisis de los meses de nacimiento de los jugadores de fútbol profesionales para ver si la edad no tomada en cuenta es influyente, esto es, para ver si el número de jugadores nacidos en la primera parte del año predominan sobre los nacidos en la última parte del año. También se hará un análisis para los entrenadores y, finalmente, también veremos si influye el hemisferio donde se nace a la hora de ser jugador de fútbol.

También se verá que al inicio de la primera división de fútbol no había una influencia significativa del mes de nacimiento a la hora de ser jugador de fútbol.

2. ANÁLISIS DEL MES DE NACIMIENTO DE LOS JUGADORES DE FUTBOL DURANTE LAS DOS ÚLTIMAS TEMPORADAS.

Voy a realizar varias tomas de datos para ver cuáles son los meses de nacimiento en los que nacieron los jugadores de fútbol en las últimas temporadas (temporadas 2014-2015 y 2013-2014).

Haciendo el recuento en cada temporada, tendremos un número de jugadores en cada mes de nacimiento. La suma del número de jugadores de los distintos meses de nacimiento dará como resultado el número de jugadores de la temporada.

El objetivo será ver si influye el mes de nacimiento a la hora de ser jugador de fútbol, esto es, si es más probable que una persona que haya nacido en unos determinados meses tenga una mayor probabilidad de convertirse en jugador de fútbol que otra que haya nacido en otros meses. En el caso de que la probabilidad de que una persona se convirtiese en jugador de fútbol profesional fuese independiente del mes de nacimiento de la persona se obtendrá un similar número de jugadores en cada mes de nacimiento. En el caso de que el factor mes de nacimiento fuese influyente se vería como el número de jugadores que se obtendría por mes de nacimiento no sería similar sino que con unos meses nacimiento habría más jugadores que con otros meses de nacimiento.

Es decir, lo que vamos a evaluar es si la gráfica de número de jugadores en función del mes de nacimiento es uniforme o no. Si la gráfica parece uniforme será que el factor mes de nacimiento no influye y si es no es uniforme será que si influye.

Para que esta uniformidad no se interprete solo de una forma visual, se va a usar el análisis de Kolmogorov-Smirnov que nos asegurara con un 99% de confianza si la distribución es uniforme o no, o lo que es lo mismo, nos asegurar con un 99% de confianza si la distribución si el mes de nacimiento influye a la hora de ser jugador de fútbol o no.

Las fechas de nacimientos de los jugadores de la primera división de fútbol de la temporada 2014-2015 se obtiene la siguiente tabla:

TABLA 2.1. PRIMERA DIVISIÓN ESPAÑOLA DE FÚTBOL TEMPORADA 2014-2015.													
	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMB	DICIEMBRE	TOTAL
Número de jugadores	49	50	48	45	59	40	36	31	32	22	15	28	455
Porcentajes (%)	10,8	11	10,5	9,9	13	8,8	7,9	6,8	7	4,8	3,3	6,2	

Una descripción de este contraste puede obtenerse en el libro: Peña, D. (2001): Fundamentos de estadística. Alianza editorial.

La prueba de Kolmogorov-Smirnov es una prueba no paramétrica (aquella cuya distribución subyacente no se ajusta a los llamados criterios paramétricos). Se trata de un contraste que compara la función de distribución teórica con la empírica.

Se consideran las siguientes hipótesis:

H0= la distribución es uniforme.

H1=la distribución no es uniforme.

La función de distribución obtenida a partir de la muestra recibe el nombre de función de distribución empírica y la función de distribución esperada (suponiendo cierta H_0), recibe el nombre de función de distribución teórica.

Se define la función de distribución teórica $F_n(x)$ donde n es el número de datos que contiene la muestra que en nuestro caso es 12. $F_{12}(x)$ es una función de probabilidad acumulada que engloba un trimestre y los anteriores, es la función de probabilidad esperada. $F(1)$ es en nuestro caso la probabilidad de que un jugador nazca en el primer mes que es de $1/12$, $F(2)$ es la probabilidad de que nazca en los 2 primeros meses que es $2/12$, $F(3)$ en los tres primeros meses que es $3/12$ y $F(4)$ en los 4 primeros meses que es $4/12$... hasta $F(12)$ que es la probabilidad de que nazcan en los 12 meses que sería 1. Esta función de distribución es la teórica, esto es, en la que suponemos que nacen un mismo número de jugadores en cada mes. Por tanto $F_1=1/12$; $F_2=2/12$; $F_3=3/12$; $F_4=4/12$... $F_{12}=12/12$

$F_0(x)$ es también una función de probabilidad acumulada que engloba un trimestre y los anteriores. Es la función de probabilidad que sacamos de la muestra, de los datos experimentales.

- Para el trimestre 1:

$F(1)=P(\text{mes}=1)= 49/455= 0,108$ (Probabilidad experimental de que un jugador nazca en el primer mes).

- Para el trimestre 2:

$F(2)= (45+50)/455=0,218$ (Probabilidad experimental de que un jugador nazca en los dos primeros trimestres).

- Para el trimestre 3:

$F(3) = (49+50+48)/455 = 0,323$ (Probabilidad experimental de que un jugador nazca en los tres primeros trimestres).

- Para el trimestre 4:

$F(4) = (49+50+48+45)/455 = 0,422$ (Probabilidad experimental de que un jugador nazca en los cuatro primeros meses).

Así hasta el mes 12 obteniendo:

$F(5)=0,552$; $F(6)=0,64$; $F(7)=0,719$; $F(8)=0,787$; $F(9)=0,857$; $F(10)=0,905$; $F(11)=0,938$; $F(12)=1$.

La medida de discrepancia en este caso se define como:

$$D_n = \max |F_n(x) - F_0(x)|$$

D_n son las diferencias entre la función de distribución empírica y teórica y el $\max D_n$ es la máxima de estas diferencias.

Solo queda comparar esta máxima discrepancia ($\max D_n$) con la obtenida en la tabla 1: $D_{crit}(\alpha, n)$. D_{crit} depende del número de datos y del % de confianza que se quiera usar. El % de confianza es $100(1 - \alpha)$. En este proyecto se ha usado siempre un 99% de confianza, por tanto, $100(1 - \alpha) = 99$ de donde obtenemos $\alpha = 0.01$. En esta temporada (la temporada 2014-2015) el número de datos es 455. Por tanto, yendo a la tabla 1 (situada en la última página del proyecto) sacamos en valor $D_{crit}(0.01, 455)$.

$\max D_n$ es la discrepancia máxima, cuya distribución se ha tabulado. Si la máxima discrepancia $\max D_n$ es mayor que la

encontrada en la tabla 1(Dcrítico), rechazaremos que la distribución sea uniforme (rechazaremos la hipótesis inicial H_0 , aceptaremos H_1 asegurando con un $100(1-\alpha)\%$ de confianza que la distribución no es uniforme). Este contraste tiene el inconveniente de que la distribución D_n es solo aproximada, el contraste es conservador, tendiendo a aceptar H_0 .

Análisis de Kolmogorov-Smirnov para la primera división española en la temporada 2014-2015:

Me s	Experimen tal	Empíri ca	$D_n(\text{Experimental-Empírica})$	Máxi mo D_n	$D_{crit}(0.01;455)$
1	0,108	0,083	0,025	0,307	0,076
2	0,03	0,167	0,137		
3	0,06	0,25	0,19		
4	0,1	0,333	0,233		
5	0,15	0,417	0,267		
6	0,21	0,5	0,29		
7	0,28	0,583	0,303		
8	0,36	0,667	0,307		
9	0,45	0,75	0,3		
10	0,55	0,833	0,283		
11	0,66	0,917	0,257		
12	1	1	0		

D_n es la comparación de la distribución experimental con la teórica (la que sería si hubiera el mismo número de jugadores en cada mes de nacimiento). Nos quedamos con la máxima de estas diferencias y la comparamos con la crítica (que es la que viene en la tabla). En el proyecto se ha usado siempre un 99% de confianza ($\alpha=0.01$) y el tamaño de muestra ha sido siempre

mayor de 35 por lo que se ha usado $D_{\text{crítico}} = 1,63/\sqrt{n}$ (véase tabla 1).

Se observa como $\max D_n > D_{\text{crit}}(0.01, 455)$, por tanto, podemos asegurar con un 99% de confianza que la distribución no es uniforme.

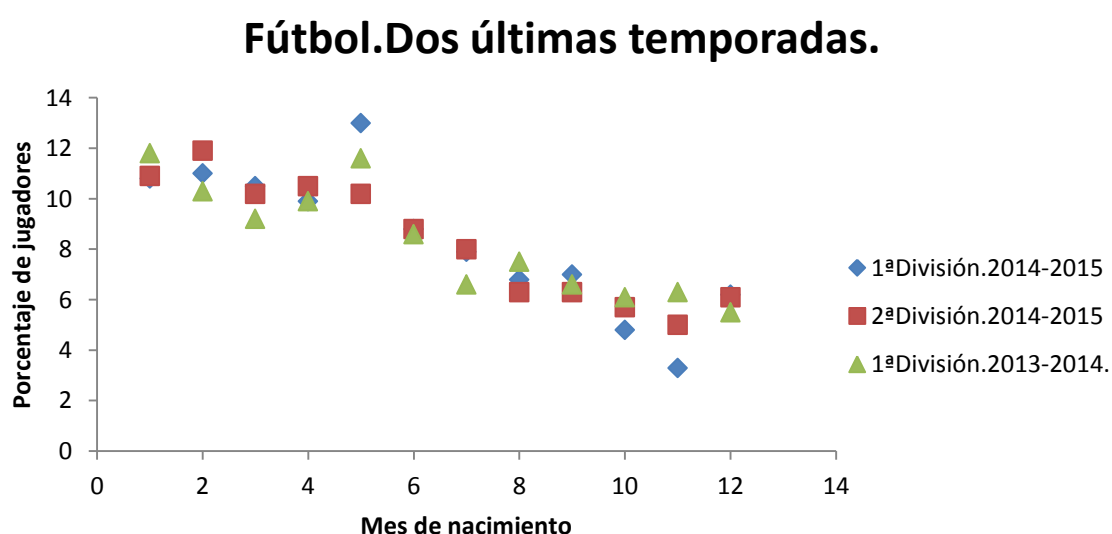
Análogamente, se procederá a analizar la influencia del mes de nacimiento de los jugadores en la segunda división de fútbol española durante la misma temporada (temporada 2014-2015) y de la primera división de fútbol española en la temporada 2013-2014. Tras realizar la toma de datos obtenemos la siguientes tablas:

TABLA 2.2. SEGUNDA DIVISIÓN ESPAÑOLA DE FÚTBOL .TEMPORADA 2014-2015.													
	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMB	DICIEMBRE	TOTAL
Número de jugadores	57	62	53	55	53	46	42	33	33	30	26	32	522
Porcentajes(%)	10,9	11,9	10,2	10,5	10,2	8,8	8	6,3	6,3	5,7	5	6,1	
TABLA 2.3. PRIMERA DIVISIÓN ESPAÑOLA DE FÚTBOL.TEMPORADA 2013-2014.													
	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMB	DICIEMBRE	TOTAL
Número de jugadores	64	56	50	54	63	47	36	41	36	33	34	30	544
Porcentajes(%)	11,8	10,3	9,2	9,9	11,6	8,6	6,6	7,5	6,6	6,1	6,3	5,5	

Haciendo una análisis de Kolmogorov-Smirnov para la segunda división de fútbol en la temporada 2014-2015 se obtiene los siguientes valores de Max Dn y D crítico: Max Dn= 0,125 y Dcrítico (0.01;522)=0,071. Realizando el mismo análisis para la primera división española en la temporada 2013-2014 se obtiene Max Dn= 0,114 y D crítico(0.01;544) = 0,07. En ambos casos se ve como Max Dn es mayor que el D crítico obtenido, por

tanto, se puede asegurar con un 99% de confianza que la distribución no es uniforme.

Con los datos obtenidos en la tabla 2.1, 2.2 y 2.3 se puede representar la gráfica del porcentaje de jugadores que nacieron en cada mes en función del mes de nacimiento para la primera y segunda división española en la temporada 2014-2015 y para la primera división española en la temporada 2013-2014 se obtiene la figura:



Se observa que existe en las dos últimas temporadas una fuerte tendencia decreciente con un claro predominio de los jugadores nacidos a primeros de año. Además el análisis de Kolmogorov-Smirnov realizado con un 99 % de confianza en los tres casos afirma que realmente existe una tendencia decreciente. La curva de la 1ª división en la temporada 2014-2015 es la que tiene la tendencia decreciente más predominante, esto se ve puesto que la diferencia entre $\text{Max } D_n$ y D crítico es mayor en este caso.

$D_{\text{crítico}}$ es parecido puesto que el nivel de confianza usado y el número total de jugadores en los tres casos son similares. Lo

que más varia es Max Dn, se puede observar que es mayor debido a que el mes con más porcentaje de jugadores y el mes con menos porcentaje de jugadores pertenecen justo a la curva de la primera división española en la temporada 2014-2015 que es la diferencia entre el mes 5(Mayo) y el mes 11(Noviembre). En esta curva, se observa una gran diferencia entre el número de jugadores que nacieron en el primer semestre y los nacidos en el segundo. También se observa como este efecto de la edad relativa no solo está en la 1ª División, sino que también en la 2ª División se da el mismo efecto.

La conclusión a la que se llega es que *en las últimas temporadas existe un fuerte efecto de la edad relativa con un predominio muy marcado de los jugadores nacidos a primeros de año.*

3. JUGADORES DE LA PRIMERA DIVISIÓN ESPAÑOLA DESDE LA TEMPORADA 1928-1929 HASTA LA TEMPORADA 2013-2014.

Ahora usaremos un número de datos mucho mayor. El objetivo es ver la evolución temporal (desde la temporada 1928-1929 hasta la temporada 2013-2014) de esta variación del número de jugadores por mes de nacimiento. ¿Siempre ha existido esta variación? ¿Cómo ha evolucionado en el tiempo? Otro de los objetivos será ver si existe una variación entre los jugadores nacidos en el hemisferio sur y los nacidos en el hemisferio norte. El objetivo es ver si el número de jugadores por mes de nacimiento es independiente de factores climatológicos, estacionales,...

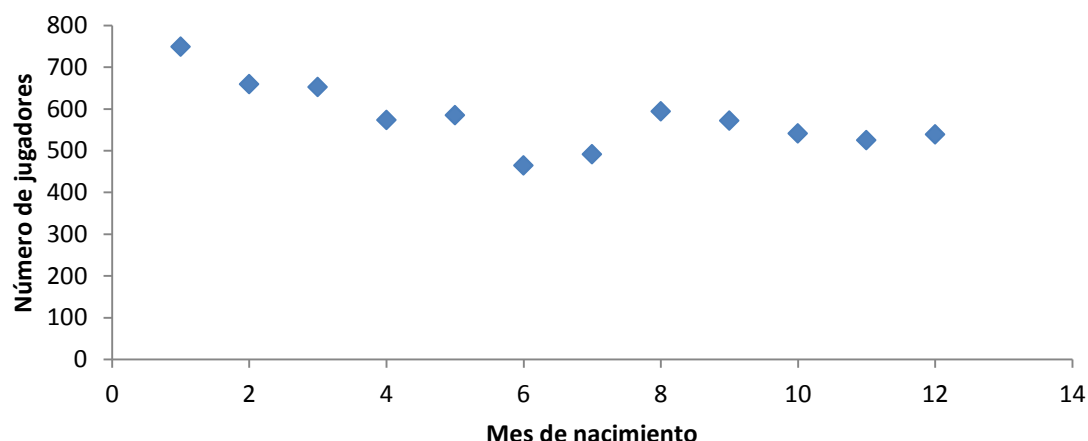
3.1. Jugadores de la primera división española nacidos en España desde la temporada 1928-1929 hasta la temporada 2013-2014.

En la fantástica web www.bdfutbol.com vienen las fechas de nacimiento de todos los jugadores de fútbol de la primera división española desde la temporada 1928-1929 hasta la temporada 2013-2014. Tras dissociar el mes de nacimiento de dichas tablas de fechas de nacimientos y contar el número de jugadores que había en cada mes obtuve la siguiente tabla:

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	TOTAL
Número de jugadores	749	659	652	573	585	464	491	594	572	541	525	539	6944
Porcentajes(%)	10,8	9,5	9,4	8,2	8,4	6,7	7,1	8,5	8,2	7,8	7,6	7,8	

Perfectamente se pueden representar los resultados:

Figura 3.1. Fútbol.Jugadores de primera división española nacidos en España desde la temporada 1928-1929 hasta la temporada 2013-2014.



Se observa como la variación es menor. La gráfica parece más uniforme.

Tras efectuar el análisis de Kolmogorov-Smirnov se obtiene máximo $D_n = 0,047$ y $D_{\text{crítico}}(0.01;6944) = 0,02$. Como $\max D_n > D_{\text{crit}}(0.01,6944)$ se puede asegurar con un 99% de confianza que la distribución no es uniforme. El número de jugadores decrece con el mes de nacimiento. Vemos como la diferencia entre $\max D_n$ y D_{crit} es muy grande a pesar de que parezca menos uniforme que las gráficas de las últimas temporadas. Esto se debe a que el número de datos que ahora se manejan es mucho mayor, por tanto, se puede rechazar la uniformidad con un alto nivel de confianza.

Relacionando esta gráfica con las de las últimas temporadas, se observa como en las últimas temporadas la tendencia decreciente era más fuerte que en la gráfica 2.1 (la de todas las temporadas). Lo que daría que pensar que podría existir una variación de la tendencia decreciente a medida que cambiamos de temporada. Es decir, que de la fuerte tendencia decreciente de las últimas temporadas se podría pasar a una temporada uniforme en las primeras temporadas para que así la tendencia de la suma de todas las temporadas (Figura 2.1) diese ligeramente uniforme. Esto se analizará con más detalle en el apartado 2.3 donde se asociara a los jugadores en tres grupos: nacidos antes de 1940, entre 1940 y 1965 y después de 1965.

3.2. Jugadores de la primera división nacidos en España y nacidos en Argentina, Brasil, Uruguay, Paraguay y Chile.

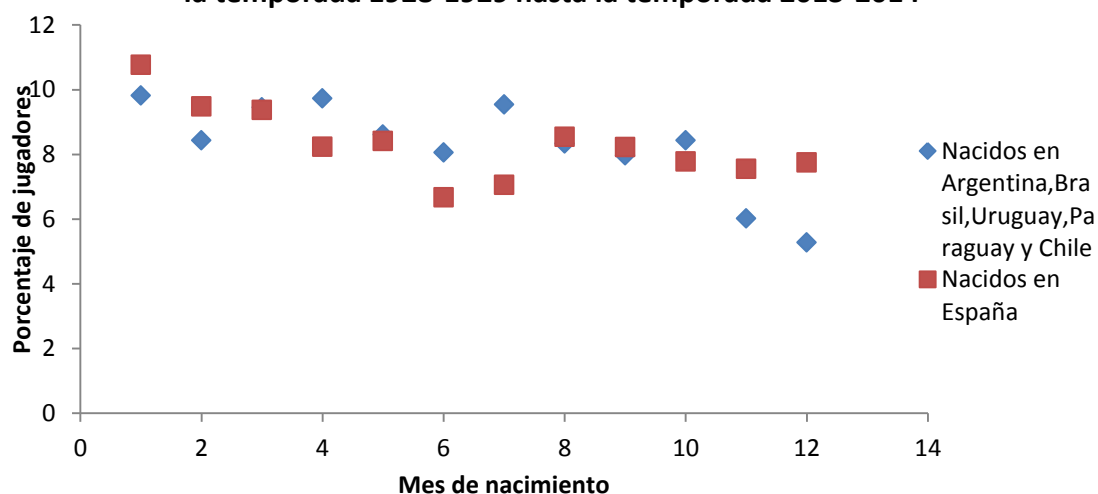
En la anterior gráfica todos los jugadores eran nacidos en España. Sin embargo, en la primera división española ha habido jugadores extranjeros tanto nacidos en países del hemisferio norte como del hemisferio sur.

Para ver esta diferencia entre hemisferios vamos a comparar a los países más representativos de ambos hemisferios pues en los países con menos números de jugadores faltan bastantes fechas de nacimiento en la base de datos. Compararemos pues los nacidos en España con los nacidos en Argentina, Brasil, Uruguay, Paraguay y Chile.

TABLA 3.2. FÚTBOL. JUGADORES DE LA PRIMERA DIVISIÓN ESPAÑOLA DESDE LA TEMPORADA 1928-1929 HASTA LA TEMPORADA 2013-2014														
HEMISFERIO		ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	TOTAL
SUR	Número de jugadores	106	91	102	105	93	87	103	90	86	91	65	57	1076
	Porcentajes(%)	9,9	8,5	9,5	9,8	8,6	8,1	9,6	8,4	8	8,5	6	5,3	
NORTE	Número de jugadores	749	659	652	573	585	464	491	594	572	541	525	539	6405
	Porcentajes(%)	10,8	9,5	9,4	8,2	8,4	6,7	7,1	8,5	8,2	7,8	7,6	7,8	

Para comparar mejor la variación del mes de nacimiento entre ambos hemisferios, representaremos los porcentajes:

Figura 2.2. Fútbol. Jugadores de primera división española desde la temporada 1928-1929 hasta la temporada 2013-2014



El análisis de Kolmogorov-Smirnov de los nacidos en España ya está realizado en el apartado anterior. Por lo tanto haremos el análisis para los jugadores de los países del hemisferio sur en el que de nuevo, $\max D_n > D_{crit}(0.01, 1019)$, por tanto, podemos asegurar con un 99% de confianza que la distribución no es uniforme. El número de jugadores decrece con el mes de nacimiento. Vemos que la diferencia entre $\max D_n$ y D_{crit} es bastante parecida sin embargo en la gráfica la variación parece ser muy similar y no depender del hemisferio. Esto es debido a que el número de jugadores nacidos en España es de 6944 y el número de los nacidos en Argentina, Brasil, Uruguay, Paraguay y Chile es 1019.

¿Hay diferencias entre ambos hemisferios?

Como se acaba de concluir, parece que no y a la vista de las gráficas cualquiera podría pensar que las diferencias son

pequeñas. De todas formas se puede comparar ambas usando el análisis de Kolmogorov-Smirnov, con la diferencia de que ahora la función de distribución teórica será la correspondiente a los jugadores nacidos en España y la experimental será la correspondiente a los jugadores nacidos en Argentina, Brasil, Uruguay, Paraguay y Chile.

Aquí están los resultados del análisis:

Mes	Experimental(Hemisferio Sur)	Teórica(Hemisferio norte)	Dn	Max Dn	Dcrítico.
1	0,099	0,108	0,009	0,043	0,05
2	0,184	0,203	0,019		
3	0,279	0,297	0,018		
4	0,377	0,379	0,002		
5	0,463	0,463	0		
6	0,544	0,53	0,014		
7	0,64	0,601	0,039		
8	0,724	0,686	0,038		
9	0,804	0,768	0,036		
10	0,889	0,846	0,043		
11	0,949	0,922	0,027		
12	1	1	0		

Se puede observar como $\max D_n < D_{crit}(1019, 0.01)$, por tanto, podemos asegurar con un 99% de confianza que no hay diferencias entre dichas distribuciones. Es decir la tendencia es independiente del hemisferio en el que esté situado el país. Las personas nacidas en Sudamérica se enfrentan a la misma influencia del mes nacimiento a la hora de convertirse en jugadores de fútbol profesional que los nacidos en España. El

factor mes de nacimiento es el mismo, no depende del hemisferio.

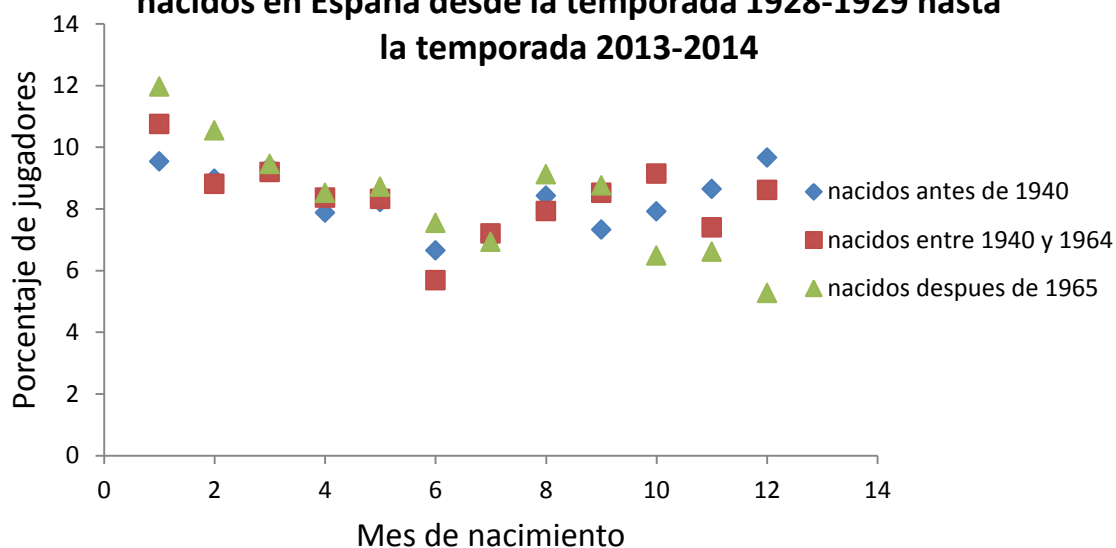
3.3 Jugadores de fútbol nacidos en España de la primera división española disociados por su año de nacimiento en tres grupos: nacidos antes de 1940, nacidos entre 1940 y 1965 y nacidos después de 1965.

Ahora vamos a disociar a los jugadores que estuvieron en la primera división española y que nacieron en España en tres grupos: nacidos antes de 1940, nacidos entre 1940 y 1965 y nacidos a partir de 1965. El objetivo será ver si en los distintos periodos no se ha cumplido la uniformidad y comparar los tres grupos.

TABLA 3.3. FÚTBOL. JUGADORES DE LA PRIMERA DIVISIÓN ESPAÑOLA NACIDOS EN ESPAÑA DESDE LA TEMPORADA 1928-1929 HASTA LA TEMPORADA 2013-2014														
NACIDOS		ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	TOTAL
Antes de 1939	Número de jugadores	225	212	221	186	194	157	165	199	173	187	204	228	2123
	Porcentajes	9,5	9	9,4	7,9	8,2	6,7	7	8,4	7,3	7,9	8,6	9,7	
Entre 1940 y 1964	Número de jugadores	221	181	189	172	171	117	148	163	175	188	152	177	2054
	Porcentajes	10,8	8,8	9,2	8,4	8,3	5,7	7,2	7,9	8,5	9,2	7,4	8,6	
Después de 1965	Número de jugadores	295	260	233	210	215	186	171	225	216	160	163	130	2464
	Porcentajes	12	10,6	9,5	8,5	8,7	7,5	6,9	9,1	8,8	6,5	6,6	5,3	

Se representan los porcentajes para ver mejor las diferencias de los tres grupos.

Figura3.3.Fútbol.Jugadores de primera división española nacidos en España desde la temporada 1928-1929 hasta la temporada 2013-2014



En la gráfica de arriba los puntos en azul corresponde a los nacidos antes de 1940, los rojos a los nacidos entre 1940 y 1965 y el verde a los nacidos después de 1965.

Haremos un análisis de Kolmogorov-Smirnov en cada grupo por separado obteniendo para cada uno de los tres casos un valor de la máxima diferencia, $\max D_n$, y un valor de D crítico.

Nacidos	antes de 1940	entre 1940 y 1965	después de 1965
Max D_n	0,029	0,039	0,076
$D_{crítico}$	0,034	0,036	0,033
Tendencia	Uniforme	Decreciente	Decreciente

Los valores de $D_{crítico}$ son similares debido a que el número total de datos en cada uno de los casos de la tabla 2.3 son similares y el nivel de confianza usado es el mismo.

En cuanto a los valores de $\text{Max } D_n$ ya se veía en la figura 2.3 que el que iba a tener la mayor diferencia eran los nacidos después de 1965; observamos que las diferencias entre los nacidos en diciembre y enero es la diferencia más grande de la gráfica de tres curvas. Si el valor máximo de la diferencia supera al valor crítico entonces la tendencia será decreciente y si el valor máximo de la discrepancia no consigue superar el valor crítico la tendencia se considerara uniforme con un 99 % de confianza.

En nuestro caso vemos que para los nacidos antes de 1940, la distribución era uniforme, es decir, no existía un efecto de la edad relativa. A partir de 1940 empieza a existir dicho efecto de la edad relativa que se va haciendo cada vez más marcado como lo muestra el análisis de Kolmogorov-Smirnov, donde pasamos de la uniformidad (antes de 1940) a una tendencia decreciente poco marcada (entre 1940 y 1965) y de ahí a una tendencia aún más decreciente (después de 1965). *Por tanto, la tendencia decreciente no estuvo siempre ahí, no hubo siempre un efecto de la edad relativa y además ese efecto de la edad relativa aumenta conforme avanzamos hacia las últimas temporadas.*

3.3.1 Conclusiones acerca de la evolución temporal.

Se ha comprobado que para los jugadores nacidos antes de 1940 (entre 1900 y 1940) la distribución era uniforme. Incluso en este grupo hemos observado que el mes de nacimiento con más jugadores era diciembre; sin embargo, entre 1940 y 1965 la distribución deja de ser uniforme a tener más jugadores nacidos en la primera parte del año. A partir de 1965 esta tendencia decreciente se hace aún más marcada, habiendo una mayor

diferencia entre el número de jugadores que nacen a principios de año y el número de jugadores que nace a finales.

Incluso, hemos analizado por separado las últimas temporadas donde este decrecimiento esta exageradamente marcada.

Por tanto, la evolución del fútbol viene marcada por un crecimiento del número de jugadores nacidos a principios de año y una disminución del número de jugadores nacidos a finales de año.

En algunos trabajos anteriores sobre edad relativa, se ha intuido que la causa de esta no uniformidad es debida a una falta de motivación o de capacidad de liderazgo, debido a que en las escuelas a edad temprana los mayores de la clase se posicionan como líderes y que esto aumenta la capacidad de liderazgo en la edad adulta. Sin embargo, cuando se observa que antes de 1940 la distribución era uniforme es difícil estar de acuerdo con esta causa.

A medida que el fútbol fue cogiendo más relevancia se empezó a formar a los jugadores. Los mejores jugadores son seleccionados a una determinada edad para ser formados aumentando sus habilidades, generalmente, en mayor medida de lo que las aumentarían si no la recibieran. De manera que la formación se convierte en parte del negocio y, generalmente, a mayor tiempo de formación mejor es el resultado. Esto implica realizar la selección de entrada a la formación a una edad cada vez más temprana.

A medida que la selección se hace a una edad más temprana, el mes de nacimiento cobra más importancia pues las

diferencias de desarrollo entre personas de un mismo curso académico son más grandes.

Es decir, se ha llegado a la conclusión de que dicha diferencia no se debe a una mejor capacidad de liderazgo o de motivación, sino a una selección a una edad cada vez más temprana, favoreciendo cada vez más a jugadores nacidos a principios de año.

Al ser cada vez más influyente el mes de nacimiento o diferencia de edad no tenida en cuenta, el factor mes de nacimiento cobra cada vez más importancia. El no tenerse en cuenta hace que se cometa un error cada vez más grave, una pérdida de talentos nacidos a finales de año y una ganancia de jugadores de menos nivel favorecidos por tener en realidad una edad mayor a la de sus compañeros sin que esta sea tenida en cuenta.

Un niño de 2 años es un 100% mayor que uno de un año, un niño de 3 años es un 50% mayor que un niño de 2 años, un niño de 4 años de un 33% mayor que un niño de 3 años... una persona de n años es un $100/n-1$ % mayor que una persona de n-1 años. Es decir la variación porcentual anual es $100/n-1$.

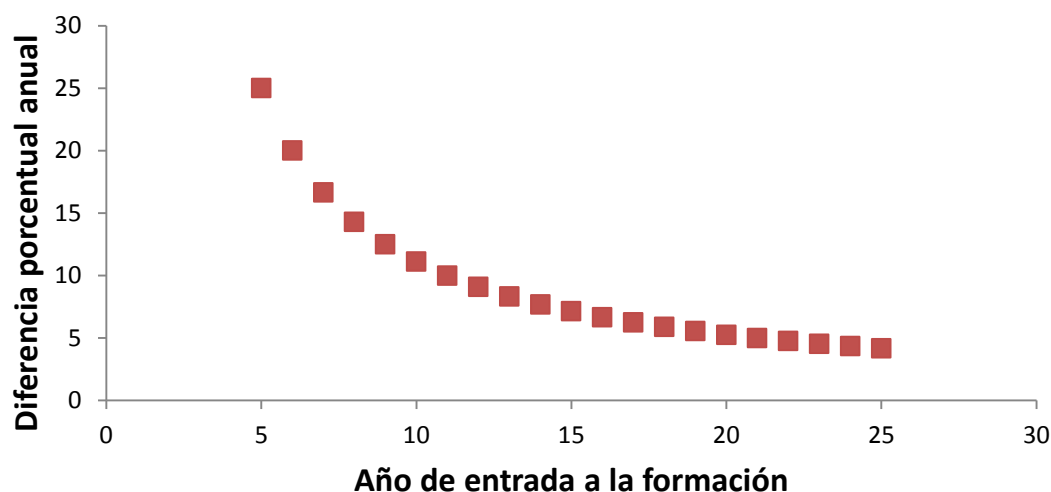
TABLA 3.3.1																		
Año	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Diferencia(%)	100	50	33,3	25	20	16,7	14,3	12,5	11,1	10	9,1	8,3	7,7	7,1	6,7	6,3	5,9	5,6

Por ejemplo, una persona con 8 años es un 14,3 % mayor que una persona de 7 años. Sin embargo una persona de 18 años es un 5,9 % mayor que una persona de 17 años.

El mismo tipo de gráfica ocurriría haciendo la diferencia porcentual por meses, horas, minutos, segundos o cualquier tipo

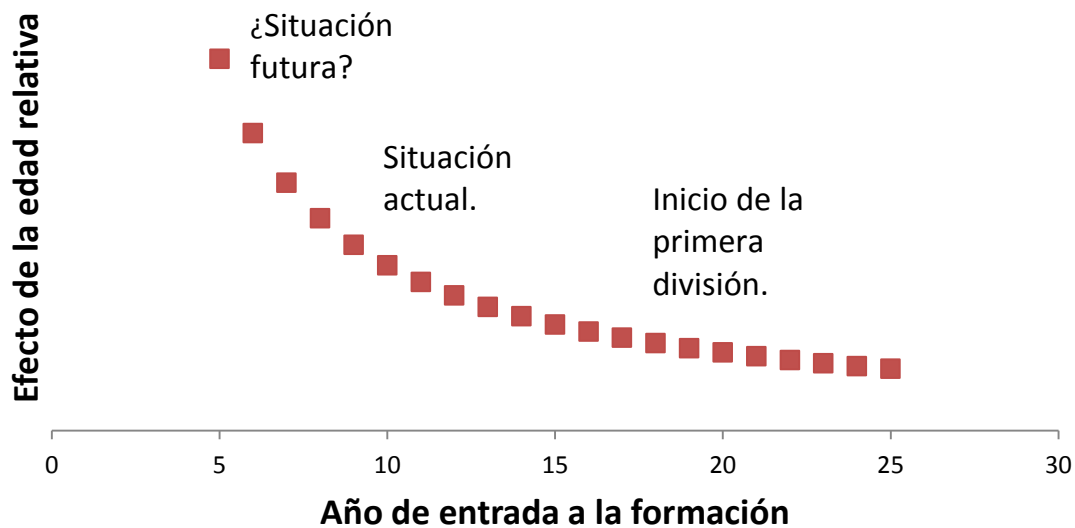
de unidad temporal que se usase para medir esta diferencia porcentual entre edades que seguirá un gráfico del estilo $1/x$.

Figura 3.3.1. $y=100/x-1$



Es el mismo tipo de exponencial ($y=1/x$) que nos hemos encontrado en las gráficas anteriores. En realidad las gráficas anteriores son como partes de esta exponencial. La parte del principio correspondería a las últimas temporadas donde la variabilidad en el mes de nacimiento era grande y la parte del final, la selección más tardía, correspondería a la parte uniforme (antes de 1940).

Figura 3.3.2.Evolución del efecto de la edad relativa



La evolución del fútbol va hacia un aumento del efecto de la edad relativa siguiendo una curva como mínimo tan dependiente del año de entrada a la formación como la de la figura 3.3.2 donde a edades de selección altas, pequeños decrecimientos en la edad de selección provocan pequeños aumentos del efecto de la edad relativa y, a edades bajas, pequeños decrecimientos de la edad de selección provocan crecimientos grandes del efecto de la edad relativa.

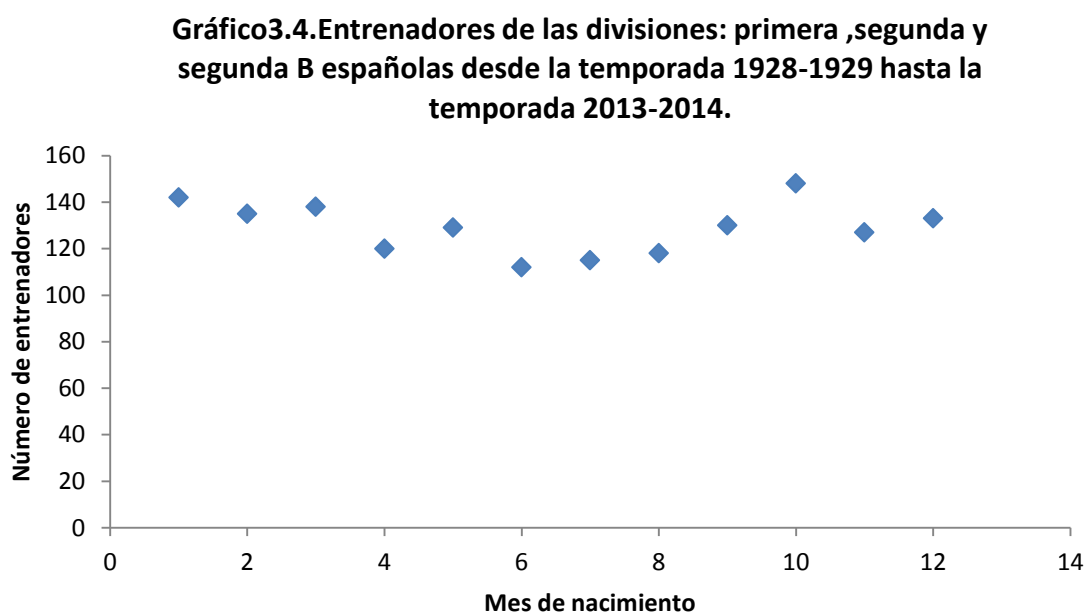
En síntesis, la evolución del fútbol va de una formación tardía a una formación temprana, de una diferencia porcentual de edad entre jugadores baja a una alta, de la uniformidad al decrecimiento del número de jugadores por mes de nacimiento, de cometer un error despreciable al seleccionar por no tener en cuenta el mes de nacimiento a un error grande por no tener en cuenta el mes de nacimiento.

3.4 Entrenadores de las divisiones: primera, segunda y segunda B españolas desde la temporada 1928-1929 hasta la temporada 2013-2014.

En la web www.bdfutbol.com disponemos también de las fechas de nacimiento de los entrenadores desde la temporada 1928-1929 hasta la temporada 2013-2014. En este caso cogeremos no solo a los de primera división sino también a los de segunda y segunda B para disponer de un número mayor de datos. Realizando el recuento obtenemos la siguiente tabla:

TABLA3.4.Fútbol. Entrenadores de las divisiones: primera, segunda y segunda B españolas desde la temporada 1928-1929 hasta la temporada 2013-2014.													
	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	TOTAL
Número de entrenadores	142	135	138	120	129	112	115	118	130	148	127	133	1547
Porcentajes(%)	9,5	9	9,2	8	8,6	7,5	7,7	7,9	8,7	9,9	8,5	8,9	100%

La gráfica es la siguiente:



Parece bastante uniforme, sin embargo, realizaremos el análisis de Kolmogorov-Smirnov para ver que conclusión arroja.

Tras realizarlo se obtiene: $\max D_n = 0,034$; $D_{\text{crítico}}(0.01;1547) = 0,041$.

Se observa como $\max D_n < D_{\text{crit}}(0.01,1547)$, por tanto, se puede asegurar con un 99% de confianza que la distribución es uniforme.

Es curioso ver como en entrenadores no se da la misma variación que en los jugadores. Teniendo en cuenta que los entrenadores han sido antes jugadores se deduce que hay un segundo efecto de la edad relativa.

Una persona nacida a principios de año es más probable que se convierta en jugador de fútbol (primer efecto de la edad relativa), y además, un jugador de fútbol nacido a principios de año tiene menos probabilidades de convertirse en entrenador que uno que nació a finales (segundo efecto de la edad relativa).

Hay un segundo efecto de la edad relativa puesto que se pasa de un predominio de los jugadores nacidos a principios de año a la uniformidad del mes de nacimiento en los entrenadores.

4. CONCLUSIONES

Como se vio en la evolución temporal, la causa del efecto de la edad relativa es la selección temprana. A medida que la edad de selección se va retrasando, el efecto de la edad relativa va aumentando. Para que haya efecto de la edad relativa (influencia del mes de nacimiento a la hora de convertirse en jugador de fútbol) debe existir selección temprana. Si la selección es tardía no existe efecto de la edad relativa.

Liga Baseball USA vs Primera División Española

En el apartado 1, selección de jugadores, se vio como en la liga de baseball donde el mes de corte para seleccionar jugadores era agosto, se tenía que agosto era el mes de moda de la liga y que iba decreciendo el número de jugadores a medida que nos acercábamos a julio que era el mes de nacimiento con menor número de jugadores.

En la liga de fútbol española la fecha de corte es enero, ese justo es el mes en que más jugadores de la liga nacieron y va disminuyendo ese número de jugadores hasta que se llega a diciembre que es el mes en el que menos jugadores nacieron. En ambos encontramos un efecto de la edad relativa en el que el mes de corte para seleccionar jugadores es el mes de moda de la liga.

*Si el mes i es el mes de corte, entonces el mes de moda es i y el mes con menos probabilidad de ser jugador $i-1$ o $i+11$.
 $P(\text{mes}=i) > P(\text{mes}=i+1) > P(\text{mes}=i+2) > \dots > P(\text{mes}=i+11)$.*

La comparativa de las dos tendencias puede verse en la siguiente tabla:

LIGA DE BASEBALL USA	PRIMERA DIVISIÓN ESPAÑOLA
1. Selección y formación temprana.	1. Selección y formación temprana.
2. Existe efecto de la edad relativa.	2. Existe efecto de la edad relativa.
3. La probabilidad de pertenecer a la liga depende del mes de nacimiento.	3. La probabilidad de pertenecer a la liga depende del mes de nacimiento.
4. El mes de corte para seleccionar jugadores es agosto.	4. El mes de corte para seleccionar jugadores es enero.
5. Una persona nacida en agosto tiene la mayor probabilidad de convertirse en jugador.	5. Una persona nacida en enero tiene la mayor probabilidad de convertirse en jugador.
6. Una persona nacida en julio tiene la menor probabilidad de convertirse en jugador.	6. Una persona nacida en diciembre tiene la menor probabilidad de convertirse en jugador.

Hemisferio norte vs Hemisferio sur.

Haciendo uso el análisis de Kolmogorov-Smirnov se demostró que las diferencias entre el hemisferio norte y el hemisferio sur

eran despreciables y que, por tanto, el efecto de la edad relativa es independiente del hemisferio.

Evolución temporal.

En las conclusiones acerca de la evolución temporal ya se habló de la causa que provoca el efecto de la edad relativa que es la selección temprana.

Cuando la selección es tardía, el efecto de la edad relativa no existe y la distribución que siguen los meses de nacimiento es uniforme.

La existencia de selección temprana provoca que la probabilidad de que una persona sea jugadora de fútbol dependa del mes de nacimiento. Se vio que la selección tardía se dio al inicio de la liga profesional y en ese inicio la probabilidad de ser jugador de fútbol era independiente del mes de nacimiento.

A medida que avanza la actividad futbolística, la edad de selección va bajando y el efecto de la edad relativa aumentando haciendo que la probabilidad de que una persona sea jugadora de fútbol dependa en gran medida de su mes de nacimiento.

El mes de nacimiento cobra en las últimas temporadas una enorme importancia y es determinante a la hora de conseguir ser jugador de fútbol profesional.

La tendencia va hacia un crecimiento del efecto de la edad relativa que sigue una evolución como la figura 3.3.1 del estilo 1/t donde en edades de selección altas (selección tardía), el bajar un poco la edad de selección provoca crecimientos pequeños del efecto de la edad relativa. En cambio para edades de selección bajas (selección temprana), decrecimientos pequeños en la edad

de selección provoca crecimientos grandes del efecto de la edad relativa.

En la siguiente tabla se comparan ambas tendencias:

Tendencia antigua	Tendencia nueva
<i>1. Inicio de la primera división española.</i>	<i>1. Desarrollo de la primera división española.</i>
<i>2. Distribución uniforme.</i>	<i>2. Distribución decreciente.</i>
<i>3. Está presente en los jugadores nacidos antes de 1940.</i>	<i>3. Está presente en los jugadores nacidos después de 1940.</i>
<i>4. Selección tardía.</i>	<i>4. Selección temprana.</i>
<i>5. Formación poco profesional.</i>	<i>5. Formación muy profesional.</i>
<i>6. Efecto de la edad relativa despreciable.</i>	<i>6. Existe efecto de la edad relativa.</i>
<i>7. La probabilidad de ser jugador de fútbol es independiente del mes de nacimiento.</i>	<i>7. La probabilidad de ser jugador de futbol depende del mes de nacimiento.</i>
<i>8. Decrecimientos pequeños de la edad de selección provocaría crecimientos pequeños del efecto de la edad relativa.</i>	<i>8. Decrecimientos pequeños de la edad de selección provocaría crecimientos grandes del efecto de la edad relativa.</i>
<i>9. Crecimiento lento del efecto de la edad relativa.</i>	<i>9. Crecimiento rápido del efecto de la edad relativa.</i>

Jugadores vs Entrenadores

Como se vio en el apartado 3.4, la distribución de los meses de nacimiento de los entrenadores es uniforme, por tanto, debe existir un segundo efecto de la edad relativa que haga que los jugadores que nacieron a finales de año tengan mayores probabilidades de convertirse en entrenadores.

Como se mencionó en el apartado 3.4:

Una persona nacida a principios de año es más probable que se convierta en jugador de fútbol (primer efecto de la edad relativa), y además, un jugador de fútbol nacido a principios de año tiene menos probabilidades de convertirse en entrenador que uno que nació a finales (segundo efecto de la edad relativa).

En la siguiente tabla se puede comparar ambos efectos:

Jugadores	Entrenadores
1. Siguen una distribución decreciente.	1. Siguen una distribución uniforme.
2. Selección temprana.	2. Selección tardía.
3. Personas nacidas a primeros de año tienen mayores probabilidades de convertirse en jugadores.	3. Jugadores nacidos a finales de año tienen mayores probabilidades de convertirse en entrenadores.

Tabla 1.

Prueba de Smirnov-Kolmogorov. Valores críticos $D_{\max}(\alpha, n)$					
Tamaño de la muestra	Nivel de significancia α				
	.20	.15	0.10	0.05	0.01
1	.900	.925	.950	.875	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.828
4	.494	.525	.564	.624	.733
5	.446	.474	.510	.565	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.410	.490
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.392
17	.250	.266	.286	.318	.381
18	.244	.259	.278	.309	.371
19	.237	.252	.272	.301	.363
20	.231	.246	.264	.294	.356
25	.210	.220	.240	.270	.320
30	.190	.200	.220	.240	.290
35	.18	.190	.201	.230	.270
≥ 35	$1.07/\sqrt{N}$	$1.14/\sqrt{N}$	$1.22/\sqrt{N}$	$1.36/\sqrt{N}$	$1.63/\sqrt{N}$

BIBLIOGRAFÍA

Spira, G. (2008): "The boys of late Summer".

Gladwell, M. (2008): Fuera de serie (outliers). Por que unas personas tienen éxito y otras no. Taurus.

Peña, D. (2001): Fundamentos de estadística. Alianza editorial.

Base de datos histórica de fútbol [en línea].
<<http://www.bdfutbol.com>> [Consulta: Noviembre 2014].

Llames, R. (1999): "Revista de psicología del deporte, volumen 8, número 1".